
Using a HHD with a HMD for Mobile AR Interaction

Rahul Budhiraja

The HIT Lab NZ
University of Canterbury
Christchurch, New Zealand
rahul.budhiraja.dark@gmail.com

Gun A. Lee

The HIT Lab NZ
University of Canterbury
Christchurch, New Zealand
gun.lee@hitlabnz.org

Mark Billinghamurst

The HIT Lab NZ
University of Canterbury
Christchurch, New Zealand
mark.billinghurst@hitlabnz.org

Abstract

Mobile Augmented Reality (AR) applications are typically deployed either on head mounted displays (HMD) or handheld displays (HHD). This paper explores novel interaction techniques for a combined HHD-HMD hybrid system that builds on the strengths of each type of device. We use the HMD for viewing AR content and a touch screen HHD for interacting with the content. A prototype system was developed and a user study was conducted comparing four interaction techniques for selection tasks.

Author Keywords

Augmented reality; wearable computer; head-mounted display; handheld display; gesture interaction

ACM Classification Keywords

H.5.1. Information interfaces and presentation: Multimedia Information Systems – Artificial, augmented and virtual realities.

Introduction

Augmented Reality (AR) is technology that overlays virtual information on the user's view of the real world in real time. One particularly promising application area is mobile AR, where portable AR systems allow the user to see virtual content superimposed over their surrounding environment wherever they are. The first



Figure 1. Using a HHD with a HMD in an AR interface.

mobile AR applications were wearable-computing systems where users would view virtual content on a head mounted display (HMD) that was connected to a backpack containing a laptop and sensors such as GPS and gyroscopes [3]. More recently, handheld displays (HHD) such as mobile phones have been used to provide large-scale access to mobile AR experiences [1]. Thus mobile AR applications have typically been delivered on either HMD or HHD, but not both.

In our research we are interested in exploring how a touch sensitive HHD can be used in conjunction with a HMD to enable new types of mobile AR interaction methods (see Figure 1). While there has been significant research on interaction methods for both HMD and HHD AR, less work has been done on systems combining both head mounted and hand held devices. This is particularly important as new lightweight HMD (e.g. Google Glass) are becoming available, and HHDs (e.g. smartphones) are widely adopted.

Related Work

The Touring Machine [3] was one of the earliest mobile AR systems that made use of a HHD in combination with a HMD. The system used a tablet computer with a stylus input for showing additional information about virtual content on the HMD. Henderson et al. [6] used a smartphone attached on user's wrist as a secondary display showing graphical user interfaces to control an AR presentation being viewed on a HMD.

Wearable computing systems use various kinds of input devices held in the hand or even attached on the user's body [9]. While HHDs could work as a pure input device [5], using a HHD in conjunction with a HMD in an AR interface also gives the possibility for multi-display

viewing. Earlier work used a desktop LCD display [4] and a projection display [7] with a HMD, so in a similar way HHDs can be also used as a secondary display for visualizing complementary information [3][6][2]. However there has been little work on methods for interaction between HMD and HHD.

Interaction between HMD and HHD

The overall focus of our work is exploring interaction methods between a HMD and a HHD. We considered two ways a HHD could be used for gesture input; (1) by making gestures on the touch screen, and (2) by moving the HHD itself. For object selection and manipulation, we adapted some existing gesture techniques from VR research, including the Sticky Finger and Head Crusher techniques [8].

In the adapted Sticky Finger method, the user moves their finger on the HHD, which causes a virtual cursor to appear in the HMD view, controlled by the finger motion on the HHD. The user then holds their finger down on the HHD around the object of interest. If the user keeps the virtual cursor on the target object for a certain period of time, the object gets selected and follows the cursor movement. In the Head Crusher method the user touches the HHD with two fingers, causing two cursors to appear, and selects the object in the HMD view by placing their fingers on the HHD so that the virtual cursors are above and below the object and holding still for a second. Once the object is selected, the user translates it by dragging either one or both of their fingers on the HHD. In this way he/she has the option of selecting the model by using the Head Crusher gesture and translating it with one finger.



Figure 2. Hardware setup of the prototype system, showing a user wearing the system, the modified HMD, and gesture input on the HHD.

Another selection method is the Tap Again gesture, where the user uses gestures on a HHD to place a cursor on the object in the HMD view, then lifts the finger briefly off the HHD and then taps down again to select the object. After being selected, as the user drags his or her finger on the HHD, the object on the HMD follows the cursor movement.

We also adapted pinch gestures common with touch screen devices for scaling virtual objects in the HMD view. The user places a pair of cursors on the side of the virtual object; By pinching outwards, the user can scale up the model and by pinching inwards, the user can scale down the model.

We were also interested in exploring gestures made with the motion of the handheld device itself. For this type of Motion Pointing interaction we used the motion of the HHD to move the cursor in the HMD view. The user holds the HHD with one hand in portrait orientation and the cursor moves towards the direction that the HHD is tilted. The speed of the cursor movement is controlled by the tilted angle and users can select an object under the cursor by touching the touch screen with their thumb.

Figure 2 shows our prototype system implemented using a monocular HMD, a Sony Vaio Ultra Mobile PC, and a Samsung Galaxy SIII Android smartphone. The HMD is a modified Vuzix Wrap 1200 with one eyepiece removed to enable to user to see the real world with a virtual overlay. The HMD has an integrated head tracker, and is connected to the UMPC for graphics generation. User input on the smart phone was captured by a native Android application and transferred to the UMPC using Wifi. The Sony VAIO was

running Windows XP and openFrameworks was used to create a simple AR application for the HMD.

Besides interaction methods for object selection and manipulation tasks, there are a wide range of other interaction techniques that could be done with combined hybrid HHD and HMD systems, such as use of cross-device gestures.

Cross-device swipe gesture

In our system, the user can select an object on the HMD and then swipe horizontally on the HHD to show additional information about it on the HHD screen. Swiping is similar to dragging an object from one display onto another and feels natural on a monocular HMD and HHD. As shown in Figure 3, to perform the cross-device gesture the user switches to the cross-device gesture mode through a menu option (A). The user then places a cursor on the object for which he wants additional details (B) and performs the swipe gesture while dragging his finger towards the end of the screen (C). Additional information about the object appears on the HHD (D).

Cross-device gestures could be also used to move objects from the HHD screen into the AR HMD view. For example, the HHD could show a list of objects as 2D icons and when one is selected and swiped off screen, it could appear in the AR view as a 3D object, enabling quick placing of objects in the AR scene.

User Study

Using our prototype we conducted a user study to investigate which interaction method was the best for object selection tasks. The four methods were Sticky Finger (SF), Head Crusher (HC), Tap Again (TA), and

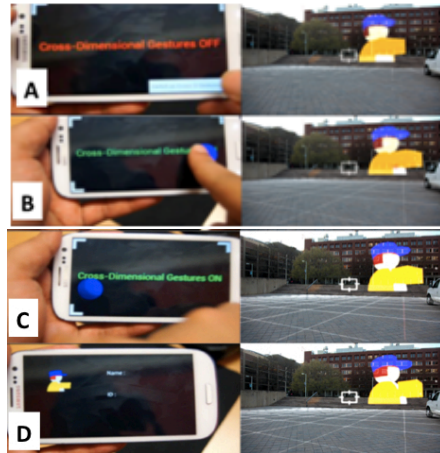


Figure 3. Cross-device swipe gesture.

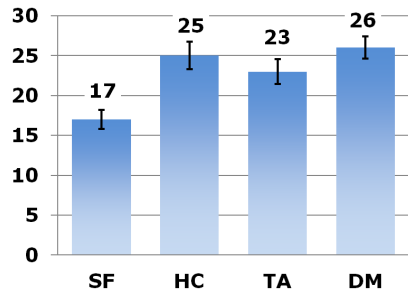


Figure 5. Results – average error (in pixels, error bar: +/- S.E.)

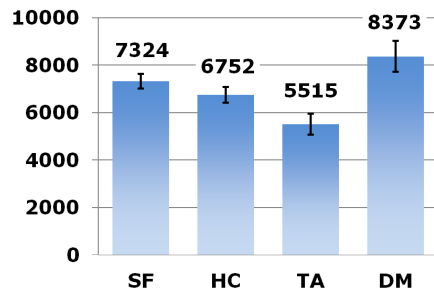


Figure 6. Results – average task completion time (in milliseconds, error bar: +/- S.E.)

Device Motion gestures (DM). For each we compared user performance in terms of accuracy and time taken, and results from a usability questionnaire.

Experimental design

Prior to the experiment, the participants were given a description of the gestures, how to perform them and how to select an object. A training phase was used where participants selected five targets with each of the gestures. Once familiar with the technology, each participant then performed a selection task using the four different methods. The study was a within group design where each participant tried all four conditions in a counterbalanced order.

The interaction task involved selecting a red bulls-eye target object in the AR scene by using one of the interaction methods (see Figure 4). The targets were randomly placed around the participant, so that he or she had to turn around to find the target and select it. The green circle is the touch position of the user's finger on the HHD mapped to the HMD. The user would select 20 of these targets one after another for each condition. A radar display in the lower left corner (see showed the location of all the targets and helped the participant complete the task.

At the end of each condition, we asked the users to fill out a usability questionnaire with eight questions and respond to an open question asking for the participant's opinion of the interaction method used. At the end of the all four conditions, we also asked the participants to fill out a questionnaire including questions asking which interaction method they preferred, we also asked them to freely describe their overall opinions and provide comments about the interaction methods.



Figure 4. HMD screen showing the experimental task. Users had to select the red target with the green circle.

Results

The study had 12 participants, all male students between the ages of 19 to 32. All participants had prior experience of using touch screen phones, and their experience of mobile AR varied from none to frequent. The participants' also had a wide range of 3D motion interface experience varying from using none of the interfaces we had listed to those who had used all of them.

For task performance, we measured error and task completion time. The error was defined as the distance from the cursor and the target object measured in pixels at the time of selection. The maximum possible error was 50 pixels; the radius of the target. Figure 5 summarizes the result. Overall SF was the technique producing the least error. A repeated measures ANOVA with a Greenhouse-Geisser correction determined that mean errors differed statistically significantly between the different gestures ($F(2.008, 22.088) = 10.411$, $p = 0.001$). Post hoc tests with Bonferroni correction revealed that there was a statistically significant

-
- Q1. I was performing well.
- Q2. The given interface was Useful for completing the task.
- Q3. The given interface was Easy to use.
- Q4. The given interface was Easy to learn.
- Q5. The given interface was Intuitive.
- Q6. The given interface was Natural.
- Q7. The given interface was Mentally stressful.
- Q8. The given interface was Physically stressful.
-

Table 1. Usability questions

	SF	HC	TA	DM	<i>p</i>
Q1	6.25	6.00	7.25	5.00	0.01
Q2	6.00	5.25	7.25	3.92	0.01
Q3	6.08	5.58	7.25	3.75	0.01
Q4	7.08	6.50	7.67	5.50	0.02
Q5	6.42	5.33	7.17	4.25	0.01
Q6	5.92	4.58	6.25	3.92	0.04
Q7	4.00	4.42	3.50	5.25	0.01
Q8	3.25	3.25	2.75	4.50	0.01

Table 2. Results – usability questions (mean of Likert scale rating 1~9)

difference between SF and the other three methods (HC: $Z=-2.831$, $p=0.005$, TA: $Z=-3.071$, $p=0.002$, DM: $Z=-3.063$, $p=0.002$). There were no other significant differences between the conditions.

Figure 6 summarizes the average task completion time in milliseconds for selecting each target. Overall Tap Again was the fastest method. A repeated measures ANOVA with a Greenhouse-Geisser correction determined that task completion time differed statistically significantly between the different methods ($F(2.041, 22.449) = 18.419$, $p < 0.001$). Post hoc tests revealed there was a statistically significant difference between TA and the other methods (SF: $Z=-2.981$, $p=.003$, HC: $Z=-2.667$, $p=.008$, DM: $Z=-3.059$, $p=.002$). There are no other significant differences between the other conditions.

The participants answered eight usability questions (see Table 1) on a Likert scale ranging from 1 (strongly disagree) to 9 (strongly agree). The mean values of the participant responses are shown in Table 2. The Friedman test showed there was a statistical difference for all questions ($p < 0.05$ for all). Post-hoc tests using Wilcoxon signed-rank test with Bonferroni correction showed significant difference between TA and DM for questions Q1 ($Z=-2.931$, $p=.003$), Q2 ($Z=-2.966$, $p=.003$), Q3 ($Z=-2.834$, $p=.005$) and Q5 ($Z=-2.756$, $p=.006$). This means that users felt that they performed better with the TA gesture than the DM gesture (Q1), that the TA interface was more useful for completing the task than the DM gestures (Q2), that the TA interface was easier to use than the DM gesture (Q3), and that the TA gesture was much more intuitive than the DM gesture (Q5).

After trying all four methods, users were asked to choose which interaction method they preferred, and to write their own comments. Overall, 8 participants (66%) preferred TA most and two (16%) preferred SF. One participant responded that both methods were equally good and one preferred the HC method. People liked that TA was fast and there was a minimal delay in selecting the target compared to SF.

Four participants (33%) explicitly mentioned that Head Crusher was not intuitive while six (50%) cited difficulties in keeping two fingers vertical and moving it along the HHD's touchscreen. Only one user chose Head Crusher as his preferred gesture, so it seems that it is not a suitable gesture for usage across a mainstream audience. While the feedback for the Sticky Finger gesture has been positive, eight (60%) of the participants felt the time they had to hold their finger still was too long.

Discussion

The user study results have shown the feasibility of using input on a HHD to select targets in an HMD view of an AR scene. However, there were significant differences in both the measured data and the users' opinions of the gestures. The significant differences in user opinion between TA and DM showed that users prefer on-screen touch gestures to using device motion based gestures. The study also showed that the choice of gesture could depend upon the type of application. For applications requiring fast selection, TA is most suitable whereas SF is more suited for applications requiring accuracy. Analysis of users' preferences showed that almost all users preferred TA for selection, making it a suitable gesture for widespread use over a large population. Users mentioned that they liked TA

the most because it was easy to learn and could be performed quickly.

Conclusion

This research investigated the usefulness of using input with a HHD for mobile AR applications viewed on a HMD. In this paper we have described several possible techniques for capturing gesture-based input with HHD and using that to manipulate the content in the HMD AR scene. These include gestures made on the HHD touch screen as well as motion of the HHD itself. In a selection task we found that the Tap Again gesture performed the fastest and was most preferred by users, while the Sticky Finger gesture was the most accurate. This suggests that for HHD-HMD hybrid systems, touch gestures on the device might be preferable to device motion based input, especially for selection tasks.

This research could be extended in a number of ways. We have only explored how HHD gestures could be used to select objects in the HMD AR view. We would like to investigate a richer range of object manipulation such as positioning, scaling, copying, deletion, etc. This would include implementing a variety of interaction techniques as well as conducting formal user studies with more users.

We would also like to further develop our efforts in designing interaction methods for other tasks. Finally, we could explore the application space that these HHD-HMD hybrid systems could be ideal for. Our work to date has been focused on user interaction without considering specific applications, so we would like to work with application developers to explore the design space.

References

- [1] Bane, R. and Hollerer T. Interactive tools for virtual x-ray vision in mobile augmented reality. In *Proc. ISMAR '04*, IEEE (2004), 231-239.
- [2] Benko, H., Ishak, E. W., and Feiner, S. Collaborative mixed reality visualization of an archaeological excavation. In *Proc. ISMAR'04*, IEEE (2004), 132-140.
- [3] Feiner, S., MacIntyre, B., Höllerer, T., and Webster, T. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Personal Technologies 1*, 4 (1997), 208-217.
- [4] Feiner, S. and Shamash, A. Hybrid user interfaces: breeding virtually bigger interfaces for physically smaller computers. In *Proc. UIST'91*, ACM (1991), 9-17.
- [5] Ha, T. and Woo, W. ARWand: Phone-based 3d object manipulation in augmented reality environment. In *Proc. ISUVR'11*, IEEE (2011), 44-47.
- [6] Henderson, S. J. and Feiner, S. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *Proc. ISMAR '09*, IEEE (2009), 135-144.
- [7] Kok-Lim, L., Ilie, A., Welch, G., and Lastra, A. Combining head-mounted and projector-based displays for surgical training. In *Proc. VR'03*, IEEE (2003), 110-117.
- [8] Pierce, J. S., Forsberg, A. S., Conway, M. J., Hong, S., Zeleznik, R. C., and Mine, M. R. Image plane interaction techniques in 3d immersive environments. In *Proc. I3D'97*, ACM (1997), 39-43.
- [9] Thomas, B., Tyerman, S., and Grimmer, K. Evaluation of three input mechanisms for wearable computers. In *Proc. ISWC '97*, IEEE (1997), 2-9.